4А

УДК 004.932

А.В. Агарков
Институт проблем искусственного интеллекта МОН Украины и НАН Украины, г. Донецк
aav@iai.donetsk.ua

Сегментация изображения на основе его описания в виде графа

В статье рассматривается проблема сегментации изображения на основе использования графа-описания. Для проведения сегментации рассматривается только граф-описание, построенный на основе использования мультимасштабного представления, полученного на основе свёртки исходного изображения с функцией Гаусса. Искомым сегментам соответствуют связные подграфы данного графа, которые получаются удалением вершин и рёбер, не удовлетворяющих заданным условиям. Рассмотрены свойства получаемых сегментов и их отношения.

Введение

Сегментация изображений до сих пор остаётся актуальной задачей из области технического зрения, поскольку используется при поиске и распознавании объектов, составлении дескрипторов и подписей изображений для больших баз данных [1], редактирования изображений и т.п. Сегментирование изображения – это разбиение его на области, содержащие пиксели, объединённые каким-либо общим свойством.

Основные подходы при сегментации – это 1) выделение границ областей и 2) выделение самих областей. При первом подходе реализуется поиск пикселей, являющихся границами искомых областей согласно какому-либо критерию [2], [3]. В частности, используются маски для нахождения пикселей, соответствующих границам между областями с различной яркостью, градиентные методы, использующие при отборе пикселей границ значения первой и второй производных значений яркости, методы из техники графов для отбора пикселей для получения оптимальных границ.

При втором – объединение пикселей согласно выбранному свойству. При этом используются методы пороговой сегментации, метод водораздела, метод разбиения и объединения областей и т.д. Также используются методы для выделения текстурированных областей. Например, широкое распространение получил метод использования минимальных сечений в графе [4], [5], позволяющий довольно эффективно выделять на изображении текстурированные области. Также для этой цели применяются методы, использующие Марковские модели [6], [7].

Все перечисленные выше методы используют в качестве начальных данных значения яркостей отдельных пикселей, образующих изображение. В отличие от этого, в настоящей работе предлагается способ сегментации изображений на основе анализа и разбиения графа, описывающего (представляющего) изображения, способ построения которого описан в работах [8-10]. Данный граф строится на основе использования структурных элементов, поиск которых основан на модификации метода выделения локальных черт [11], [12].

Целью данной работы является разработка метода сегментации изображений на основе использования графа-описания и исследование свойств полученных сегментов и их отношений.

Построение графа, описывающего изображение

Для построения графа, описывающего изображение, используется мультимасштабное представление на основе свертки функции яркостей изображения с функцией Гаусса и её производными.

Мультимасштабное представление изображения на основе применения функции Гаусса представляет собой упорядоченное множество изображений, полученных из исходного путём свёртки с ней:

$$L(x, y, \sigma_i) = G(x, y, \sigma_i) * I(x, y),$$

где $G(x, y, \sigma_i)$ – функция Гаусса с масштабом σ_i, $\sigma_i = k\sigma_{i-1}$, $(*)$ – операция свертки, $I(x, y)$ – исходное изображение. Данное множество изображений образует в объединённом трехмерном пространстве (плоскость изображения – масштаб сглаживания) дискретную функцию $L(x, y, \sigma)$. Для увеличения скорости построения мультимасштабного представления пользуются октавами исходного изображения – при увеличении масштаба сглаживания в два раза во столько же раз уменьшается исходное. Соответственно, множитель k выбирается так, чтобы $k^{N_o} = 2$, где N_o – количество масштабов в одной октаве.

Однако для выделения структурных элементов изображения используется мультимасштабное представление, построенное на основе свёртки с ядром, которое представляет собой разность функций Гаусса с различным масштабом сглаживания, известное как оператор DoG:

$$D(x, y, \sigma_i) = \left(G(x, y, \sigma_i) - G(x, y, \sigma_{i-1})\right) * I(x, y) = L(x, y, \sigma_i) - L(x, y, \sigma_{i-1}).$$

Данное представление также образует дискретную функцию $D(x, y, \sigma)$.

Локальные черты изображения выделяются из условия достижения экстремума функции $D(x, y, \sigma)$. При этом координаты (x, y) определяют центр области, соответствующей локальной черте, а значение σ – её характерный масштаб. Локальные черты, выделяемые с помощью данного оператора DoG, соответствуют пятнообразным областям (blobs) однородной яркости.

Множество выделенных локальных черт $E = \{E_j(x_j, y_j, \sigma_j)\}$ составляет основу для описания изображения и объектов, которые его составляют.

Для формирования структурных элементов, составляющих основу описания изображения, используются области изображения, которые соответствуют экстремумам по направлению в плоскостях мультимасштабного представления $D(x, y, \sigma_i)$. То есть используются экстремумы по направлению функций $D(x, y, \sigma = const)$. Данные экстремумы образовывают хребты и тальвеги данных функций (далее – характерные линии), которые, в свою очередь, соответствуют протяжённым деталям изображения. Как следует из свойств оператора DoG, такими деталями могут быть либо границы, либо протяженные области однородной яркости. Причём характер детали можно определить из поведения характерных линий с изменением масштаба σ. Если положение линии изменяется, то она соответствует границе, если положение линии остаётся стабильным – протяжённой области с однородной яркостью.

На рисунках представлены характерные линии исходного изображения, выделенные для различных масштабов сглаживания.

Предлагается следующая процедура для выделения дополнительных структурных элементов. В каждой плоскости мультимасштабного представления $D(x, y, \sigma_i)$ произ-

водится поиск множества экстремумов по направлению $T = \{p_j = p(x_j, y_j, \sigma_{i,j})\}$ функции $D(x, y, \sigma = const)$. Данное множество разбивается на непересекающиеся подмножества $\overline{T} = \{t_j : \bigcup_j t_j = T; \forall j, k,, j \neq k, t_j \cup t_k = \varnothing\}$, каждое из которых состоит из экстремумов, имеющих одинаковые координаты в плоскости изображения, и расположенные в последовательности из соседних плоскостей:

$$t = \{p(x_j, y_j, \sigma_{i,j}) : \forall j, k, j \neq k, (x_j, y_j) = (x_k, y_k),\ \{\sigma_{i,j} : i = i_1, ..., i_{N_t}, i_{m+1} = i_m + 1\}\},$$

где N_t – количество экстремумов p в множестве t. Множество $\overline{T}$ разделяется на два непересекающихся подмножества $\overline{T}_{\max}$ и $\overline{T}_{\min}$, которые образованы локальными максимумами и минимумами соответственно.

Каждое множество $t_i \subset T$ образует структурный элемент. Данный структурный элемент характеризуется координатами в плоскости изображения (x_i, y_i), масштабом $\sigma_{extr} = \arg\min_{p(x,y,\sigma)\in t_i} \max\left(D(x, y, \sigma)\right)$ (дальше – характерный масштаб) и величиной диапазона масштабов, в котором он сохраняет постоянное положение (далее – глубина диапазона). Для того чтобы глубина диапазона не зависела от масштаба, она характеризуется мощностью множества t_i. Причём следует учитывать, какого типа экстремумы составляют множество t_i – максимумы или минимумы. Если минимумы, то значение масштаба должно соответствовать минимальному значению $D(x, y, \sigma)$, если максимумы – максимальному. Также данный структурный элемент характеризуется значением $L(x, y, \sigma_{extr})$ – характерной яркостью.

Следует заметить, что $E \subset T$, поскольку экстремум функции $D(x, y, \sigma)$ также является экстремумом по направлению для $D(x, y, \sigma = const)$. Это позволяет описывать локальные черты при помощи дополнительных структурных элементов, расположенных в окрестности их центров.

Множество выделенных локальных черт $E = \{E_j(x_j, y_j, \sigma_j)\}$ составляет основу для описания изображения и объектов на нём. Описание изображения строится на основе применения графа. Пусть граф $G_I^T(V^T, E^T)$ является описанием изображения $I(x, y)$. Каждой вершине данного графа ставится в соответствие структурный элемент, соответствующий множеству t_i. Причём каждой вершине соответствует только один структурный элемент. Ребрами соединяем вершины, если соответствующие им структурные элементы являются соседями.

Соседство разделяется на четыре типа: 1) соседние структурные элементы, относящиеся к одной детали изображения, у которых центры являются соседними пикселями; 2) ближайшие структурные элементы, относящиеся к соседним деталям изображения одного типа; 3) ближайшие структурные элементы, относящиеся к соседним деталям изображения разного типа; 4) структурные элементы, центры которых расположены рядом и которые преобразуются друг в друга при изменении масштаба сглаживания.

Ребра, отражающие соседство разных типов, маркируются по-разному. Атрибуты вершин отражают свойства структурных элементов – координаты в плоскости изображения, характерные масштаб и яркость. Однако атрибут вершины также может содержать более подробную информацию о множестве t, которое соответствует данной вершине.

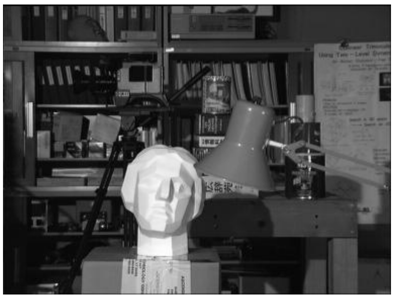


Рисунок 1 – Исходное изображение

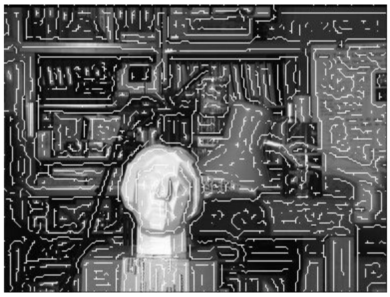


Рисунок 2 – Характерные линии, выделенные при использовании масштаба сглаживания $\sigma = 3,8$

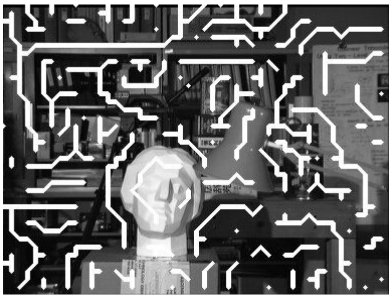


Рисунок 3 – Характерные линии, выделенные при использовании масштаба сглаживания $\sigma = 11,2$

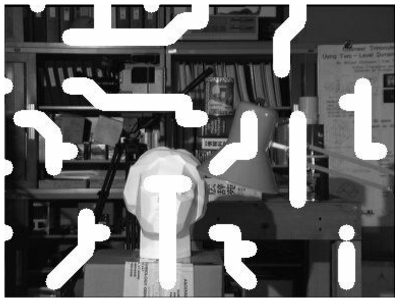


Рисунок 4 – Характерные линии, выделенные при использовании масштаба сглаживания $\sigma = 33,98$

Восстановление изображения по графу, описывающему изображение

Граф G_I^T позволяет довольно точно восстанавливать исходное изображение. Действительно, поскольку каждый структурный элемент, соответствующий вершинам данного графа, представляет собой область с однородной яркостью и центром в определённой точке плоскости изображения, то представляется возможным восстановить исходное изображение. Однако такое изображение будет лишено мелких деталей, размер которых меньше, чем начальный масштаб сглаживания.

Процедура восстановления проводится по следующему принципу – чем меньше характерный масштаб структурного элемента, тем больший он имеет приоритет при определении яркости конкретного пикселя.

Поскольку целью является исследование возможности восстановления исходного изображения по его описанию в виде графа G_I^T, то используется наиболее простой способ – в центре каждого структурного элемента рисуется окружность с радиусом, равным характерному масштабу, и характерной яркостью. Если окружности перекрываются, то яркость общих пикселей определяется характерной яркостью структурного элемента меньшего масштаба. Если масштабы равны – значение яркости в пикселе усредняется.

На рис. 5-7 показаны результаты восстановления изображения (рис. 1), описанного графом G_I^T, построенном на основе мультимасштабного представления, содержащего семь октав по семь масштабов в каждой октаве. Размер исходного изображения – 384×288. Начальное значение масштаба сглаживания – $\sigma_0 = 1,4$. Отражены

этапы восстановления, соответствующие минимальному значению характерных масштабов структурных элементов $\sigma = 22,4$ (рис. 5), $\sigma = 6,18$ (рис. 6), и конечный результат при $\sigma = 1,4$ (рис. 7).



Рисунок 5 – Восстановленное изображение, соответствующее минимальному характерному масштабу структурных элементов $\sigma = 22,4$

Как видно из рис. 5-7, чем меньше минимальный характерный масштаб структурных элементов, используемых для восстановления изображения, тем выше уровень детализации. Как и ожидалось, восстановленное изображение лишено мелких деталей, размер которых меньше минимального масштаба сглаживания. Из-за этого оно выглядит несколько сглаженным по сравнению с исходным (рис. 1). Однако визуальная оценка показывает, что результат восстановления не сильно отличается от оригинала.



Рисунок 6 – Восстановленное изображение, соответствующее минимальному характерному масштабу структурных элементов $\sigma = 6,18$



Рисунок 7 – Восстановленное изображение

Таким же образом возможно и восстановление отдельных областей изображений, которым соответствуют подграфы графа G_I^T. При этом, как и в случае со всем изображением, возможно восстановление с различной степенью детализации, которая отражена в иерархических связях между отдельными структурными элементами и областями.

Это свойство позволяет отделять распознанный объект от фона без дополнительной обработки исходного выражения. Качество восстановленного изображения также отражает адекватность предложенного описания.

Сегментация изображения

Как показано выше, областям изображения, обладающим однородной яркостью, соответствуют характерные линии, не меняющие своего расположения в плоскости изображения в достаточно большом диапазоне масштабов сглаживания. Таким образом, такой области соответствует связный подграф графа G_I^T, состоящий из вер-

шин, соединённых друг с другом рёбрами типа (1). Поскольку фактически данному подграфу соответствует характерная линия (вообще говоря, разветвляющаяся), то для её описания уместно применить аппарат скелетного представления, достаточно хорошо разработанный к настоящему времени.

Отношения соседства между областями изображения определяются из того, какие вершины соединены рёбрами типов (2) и (3). Иерархические отношения, отражающие детализацию областей, определяются рёбрами типа (4).

Таким образом, граф G_I^T позволяет для описания отдельных областей изображения использовать два подхода: первый – на основе скелетного представления, второй – на основе структуры, образуемой структурными элементами более низких уровней иерархии.

Как указано выше, выделение характерных областей изображения должно быть основано на выделении связных подграфов графа G_I^T. Поскольку основу для данного описания изображений составляют локальные черты, то каждый из подграфов, соответствующих сегментам изображения, должен содержать не менее одной локальной черты.

Предлагается в связные подграфы отбирать те вершины, отношения соседства между которыми и сами они удовлетворяют определённым условиям. Учитывая то, что каждый такой подграф должен содержать не менее одной локальной черты, процесс их выделения выглядит следующим образом – искомые подграфы инициализируются вершинами, соответствующими локальным чертам, т.е. каждая такая вершина – отдельный подграф. Затем данные подграфы расширяются за счёт вершин, смежных с теми вершинами, которые принадлежат им. Причем сами эти вершины и отношение соседства должно удовлетворять заданным условиям. Если в процессе данного расширения две смежные вершины, удовлетворяющие заданным условиям соседства, оказываются принадлежащими разным подграфам, то данные подграфы объединяются.

Фактически исходный граф разбивается на подграфы путем удаления из рассмотрения вершин и рёбер, не удовлетворяющих заданным условиям. В результате данной процедуры граф разбивается на связные подграфы. Из данных подграфов искомыми являются те, что содержат по крайней мере одну локальную черту.

В качестве параметра для отбора вершин была использована их глубина диапазона. В качестве параметров для отбора ребер – тип соседства, который они отражают, и то, как соотносятся диапазоны и характерные масштабы соседей друг с другом.

Поскольку положение структурных элементов должно быть стабильным в некотором диапазоне масштабов, то выбрасываются из рассмотрения вершины, соответствующие структурным элементам с глубиной диапазона от двух и выше.

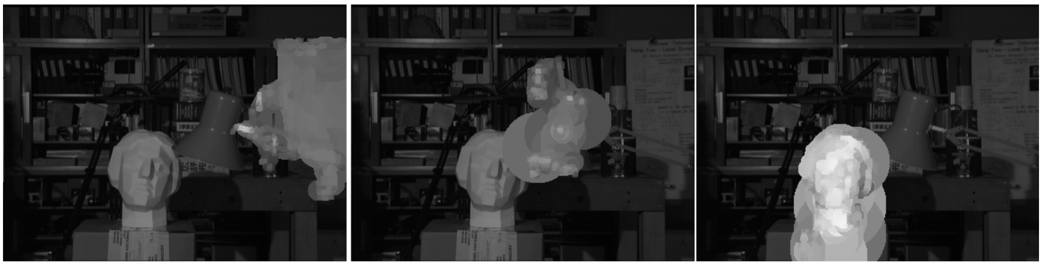
Из рассмотрения выбрасываем ребра, которые отражают отношения соседства типов (2), (3) и (4). То есть остаются только рёбра, соединяющие вершины, которым соответствуют структурные элементы, центры которых расположены в соседних пикселях. Причём характерные масштабы данных структурных элементов должны принадлежать диапазону масштабов соседа. То есть должно выполнятся условие:

$$(\sigma_{\min}^1 \le \sigma_{ext}^2 \le \sigma_{\max}^1) \wedge (\sigma_{\min}^2 \le \sigma_{ext}^1 \le \sigma_{\max}^2)\,, \qquad (1)$$

где $[\sigma_{\min}^1, \sigma_{\max}^1], [\sigma_{\min}^2, \sigma_{\max}^2]$ – диапазоны масштабов соседних структурных элементов, $\sigma_{ext}^1, \sigma_{ext}^2$ – их характерные масштабы.

Если при сегментировании (выделении подграфов) не использовать условие (1), то в результате получатся подграфы, соответствующие областям изображения, которые имеют яркость меньшую или большую по сравнению с фоном. Для того чтобы визуализировать результат сегментации, используется метод восстановления изображения на основе графа-описания, описанный выше. В качестве исходных

данных используется не весь граф, а выделенный подграф. На рис. 8 представлены примеры сегментов, выделенных на исходном изображении (рис. 1). Отдельные сегменты восстановлены на основе соответствующих им подграфов на фоне исходного изображения. Для лучшего выделения данных сегментов на окружающем фоне, исходное изображение затемнено. Как видно из рис. 8, происходит выделение областей с однородной яркостью по отношению к окружающему фону. Зачастую таким областям соответствуют отдельные объекты, которые составляют сцену на изображении.



a) b) c)

Рисунок 8 – Примеры выделенных сегментов

Как видно из рис. 8, представленные примеры выделенных восстановленных изображений, соответствующих отдельным подграфам, лишены детализации. Это происходит потому, что сами эти сегменты изображения содержат более мелкие детали, которые по яркости выделяются на окружающем фоне. Данным деталям также соответствуют подграфы/сегменты. На рис. 9 представлены примеры таких сегментов, являющихся частью большего сегмента (рис. 8c). Таким образом, данная процедура сегментации на основе применения графа-описания позволяет не только разбивать изображение на отдельные сегменты, но и выделять их составные части, соответствующие более высокому уровню детализации. Данное свойство позволяет устанавливать между сегментами иерархические отношения, отражающие степень детализации при сегментировании исходного изображения. В свою очередь, данные отношения отражают иерархическую структуру как выделенных сегментов, так и самого изображения. Данный результат есть следствие свойств графа-описания G_I^T, отражающего иерархические отношения между структурными элементами изображения.



a) b) c)

Рисунок 9 – Примеры сегментов, являющихся частью сегмента, представленного на рис. 8 c), при более высоком уровне детализации изображения

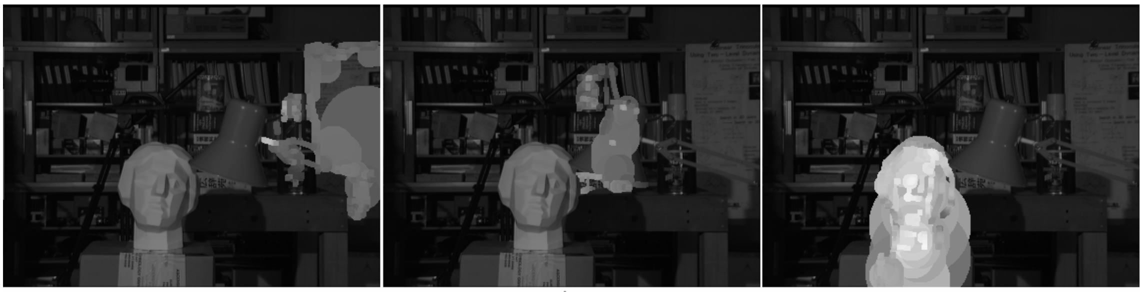
Обозначим множество подграфов, выделенных с помощью вышеописанного способа, $D^1 = \{d_i^1\}$, где d_i^1 – подграф разбиения.

Если при разбиении на подграфы использовать все вышеуказанные условия (в том числе и (1)), то результатом будут сегменты изображения, соответствующие линиям однородной яркости, толщина которых меняется плавно. Обозначим множество данных подграфов $D^2 = \{d_i^2\}$, где d_i^2 – подграф разбиения.

Поскольку способы получения множеств D^1 и D^2 отличаются лишь добавлением дополнительного условия, то очевидно, что элементы d_i^2 являются подграфами d_i^1. В частности, это означает, что каждый сегмент, соответствующий подграфам из D^1, может быть разбит на подсегменты, соответствующие подграфам из D^2.

На рис. 10 представлены подмножества сегментов, принадлежащих D^2 и соответствующих отдельным сегментам из D^1, которые представлены на рис. 8. Таким образом, элементы множества D^2 определяют внутреннюю структуру сегментов, соответствующих элементам множества D^1. Таким образом, между элементами множеств D^1 и D^2 устанавливаются иерархические отношения, определяющие внутреннюю структуру сегментов из D^1. То есть разбиение сегментов из D^1 на подсегменты при том же уровне детализации.

Таким образом, иерархические отношения между элементами D^1 отражают структуру сегментов и самого изображения при различных уровня детализации, а между элементами из D^1 и D^2 – при одном уровне детализации, соответствующем элементам из D^1.



a) b) c)

Рисунок 10 – Примеры подмножеств сегментов, принадлежащих множеству D^2, которые соответствуют элементам из D^1 и представлены на рис. 8

Отношения пространственного соседства для сегментов, соответствующих элементам множеств D^1 и D^2, определяются на основе отношений соседства между структурными элементами, составляющими основу для графа-описания изображения.

Выводы

Граф-описание позволяет проводить сегментацию изображения на основе разбиения его на связные подграфы. Для проведения данного разбиения были использованы довольно простые и логичные правила, вытекающие из свойств структурных элементов, составляющих основу для графа-описания. Данное разбиение адекватно выделяет области с однородной яркостью по сравнению с окружающим фоном. Причём выделенные сегменты изображения соответствуют различным уровням детализации и структурирования, что отражается в иерархических отношениях двух типов между ними.

Данная сегментация вполне адекватно отображает структуру изображения, несмотря на довольно простые правила, применённые для этого. Это позволяет предположить, что использование более сложных способов сегментации позволит более точно выделять сегменты изображений, соответствующие отдельным объектам и их составным частям. То есть использование графа-описания для сегментации изображений представляется довольно перспективным.

Следует также заметить, что элементы множеств D^1 и D^2 и отношения между ними составляют основу для более обобщённого описания изображения, чем у исходного графа-описания. Данное описание также удобно и естественно представить в виде графа, вершинам которого соответствуют элементы множеств D^1 и D^2, а рёбра отражают иерархические и пространственные отношения между соответствующими сегментами. Природа сегментов, соответствующих элементам множеств D^1 и D^2, и отношения между ними позволяют предположить, что данный граф возможно использовать для выделения текстур на изображении, а также в качестве обобщённого описания изображений для индексирования в базе графических данных.

Дальнейшие исследования по данной теме будут посвящены поиску способов разбиения графа-описания на подграфы, которые бы соответствовали сегментам, устойчиво и однозначно выделяемым на одинаковых объектах, независимо от изображения. Это позволит использовать выделенные сегменты для описания и распознавания различных объектов.

Литература

1. Lavrenko Jeon V. Automatic image annotation and retrieval using cross-media relevance models / Jeon V. Lavrenko, R. Manmatha // Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval. – Toronto, Canada, 2003. – P. 119-126.
2. Sumengen B. Multi-scale edge detection and image segmentation / B. Sumengen, B.S. Manjunath // Proc. European Signal Processing Conference (EUSIPCO). – Sep. 2005. – VRL ID: 146.
3. Нюнькін К.М. Виявлення границь людських облич на зображеннях // Мат. IV Всеукраїнської конф. «Оброблення сигналів і зображень та розпізнавання образів». – 1998. – С. 99-100.
4. Veksler O. Image segmentation by nested cuts / O. Veksler // IEEE Conference on Computer Vision and Pattern Recognition. – Vol. 1. – P. 339-344.
5. Yuri Boykov. Gareth Funka-Lea, Graph Cuts and Efficient N-D Image Segmentation / Yuri Boykov // International Journal of Computer Vision. – 2006. – Vol. 70, Issue 2. – P. 109-131.
6. Zhuowen Tu. Image segmentation by data-driven Markov Chain Monte Carlo / Zhuowen Tu, Song-Chun Zhu // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – Vol. 24, Issue 5. – P. 657-673.
7. Ковтун И.В. Текстурная сегментация изображений на основании марковских случайных полей / И.В. Ковтун // УСиМ. – 2003. – № 4. – С. 46-55.
8. Агарков А.В. Ієрархічне представлення зображень для застосування при вирішенні широкого кола проблем / А.В. Агарков // Оброблення сигналів і зображень та розпізнавання образів : Восьма всеукраїнська міжнародна конференція. – Київ, 2006. – С. 119-122.
9. Агарков А.В. Иерархическое представление изображения с помощью графа / А.В. Агарков // Искусственный интеллект. – 2007. – № 1. – С. 101-109.
10. Агарков А.В. Структурное описание изображений в виде графа для решения задач распознавания / А.В. Агарков // Бионика интеллекта. – 2009. – № 1(70). – С. 95-101.
11. David G. Lowe. Local Feature View Clustering for3D Object Recognition / David G. Lowe // Proc. CVPR. – 2001. – P. 682-688.
12. David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints / David G. Lowe // International Journal of Computer Vision. – 2004. – Vol. 2, № 60. – P. 91-110.
13. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – Москва : Техносфера, 2006. – 1072 с.

A.V. Agarkov
Image Segmentation Based on Using Graph-descriptor
The paper is devoted to the problem of image segmentation by graph-descriptor. Only graph-descriptor is used for segmentation. This one is built on the base of multiscale image representation which is received by using function of Gauss. Required segments corresponds to graph-descriptor subgraphs which is received by deleting of vertex and edges which does not satisfy to given conditions. The properties of segments and their relations are considered.

Статья поступила в редакцию 18.06.2010.